ERDC TN-SWWRP-06-11
November 2006

Parameter Estimation Tools for Hydrologic and Hydraulic Simulations

By Jackie P. Hallberg

PURPOSE: The purpose of this System-Wide Water Resources Program (SWWRP) technical note is to highlight some of the parameter estimation packages available and to demonstrate the application of some of those packages with U.S. Army Engineer Research and Development Center (ERDC) numerical models. Advantages and limitations of the various packages will be discussed, and examples with ADaptive Hydrology/Hydraulics (ADH) will be given.

BACKGROUND: The needs of U.S. Army Corps of Engineers users are shifting from hydrologic predictions to simulations that can be assigned a degree of uncertainty to aid in risk-analysis. Uncertainty in model predictions has several components, including modeled processes, numerical error, driving data (such as meteorological, flow, or stage data), and model input parameters. Assigning uncertainty to model predictions requires an understanding of the model's sensitivity to its input. A formal parameter estimation process gives both the best estimate of input properties for the available observations, and a quantitative measure of the model sensitivity to input. Several packages exist that are capable of parameter estimation through automated processes. However, most of these packages require numerous model calls to the simulation program to determine the best fit for the parameters. Techniques of this nature are not practical for large-scale, multiphysics simulations that are computationally intensive, such as those addressed at ERDC.

OPTIMIZATION PACKAGES: A major component of assigning uncertainty to aid in risk-analysis is determining the uncertainty related to the input parameters. Sensitivity analysis can provide useful information to this end. The standard process for sensitivity analysis is to adjust input parameters manually and monitor the changes to the model's output. This process, however, is limited to just a few adjustments and may not capture the true effect of the adjustments on the output. A more thorough method for determining the sensitivity of input changes to model output is through the use of optimization techniques.

The basic principle behind optimization is to minimize (or maximize) some objective function subject to certain constraints with respect to a vector of adjustable parameters. For instance, in surface-water problems one may wish to match cross-sectional depths by adjusting the roughness values or the elevations of marsh areas where such information is hard, if not impossible, to obtain. The engineer's desire to obtain the best design of a system lends itself to the use of numerical optimization because it offers a logical approach to design automation.

Report Documentation Page			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
NOV 2006	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Parameter Estimation Tools for Hydrologic and Hydraulic Simulations	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army Engineer Research and Development Center Vicksburg, MS 39180**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**SAR**	**13**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Some advantages of using numerical optimization are:

a. Reduction in design time.
b. A systematized logical design procedure is provided.
c. A wide variety of design variables and constraints can be dealt with.
d. Some design improvement is virtually always yielded.
e. Bias by intuition or experience in engineering is avoided.
f. A minimal amount of human-machine interaction is required.

Some limitations of numerical optimization are:

a. Computational time increases as the number of design variables increases.
b. No stored experience or intuition on which to draw is available.
c. Results could be misleading if the analysis program is not theoretically precise.
d. Global optimum design is not guaranteed.
e. Discontinuous functions are hard to deal with.
f. Significant reprogramming of analysis codes may be required for implementation (Vanderplaats 2001).

Such mathematical programming techniques have been the focus of much research.

The field of optimization can be divided into several sections, some of which overlap. Methods can be categorized into local methods and global methods, for instance, or deterministic methods versus stochastic methods. Local methods refer to those methods that are used to find local minima, i.e., the point at which the objective function is smaller than all other feasible points in the vicinity, and generally include gradient-based methods. These gradient-based methods are sensitive to initial conditions, prone to local minima, offer little reliable parameter uncertainty, and can be problematic for a large number of parameters. These methods assume information is available on the gradient vector associated with the objective function. That is, it is assumed that the gradient of the objective function with respect to the parameters being optimized can be obtained.

Global methods, or recursive methods, can be used when the gradient information is not readily available. These algorithms do not depend on a direct gradient; rather approximations to the gradients are formed from measurements of the objective function. These methods determine the global minimum, i.e., the best of all minima in the domain, by using methods such as the random sampling methods. These random sampling methods are capable of moving out of local minima and finding the global minimum. However, these methods often require too many optimization iterations and are computationally inefficient. (Fletcher 1987; Kelley 1999; Onvubiko 2000)

Deterministic methods refer to those methods that are applied to models that are fully specified. Models that offer information about the gradients, the model parameters, and the model output fall into this category. Stochastic methods are used when models cannot be fully specified because of unknowns. The uncertainty in the model is quantified to produce solutions that optimize the expected performance of the model. (Nocedal and Wright 1999)

Optimization techniques are popular as engineering aids in design and calibration of models. Watershed models have been combined with gradient-based methods such as Parameter

ESTimation (PEST) to solve multicomponent objective function parameter estimation problems by varying proportions of the flow components. The results of such methods, however, are dependent on reasonably accurate estimates of each component of the objective function (e.g., Gutierrez-Maness and McCuen 2005). Other studies combine global methods such as System Optimization and Design Algorithm (SODA) (Vrugt et al. 2005) and Generalizaed Likelihood Uncertainty Estimation (GLUE) (Beven and Freer 2001) with watershed models to calibrate models of various catchments including the River Morava in the Czech Republic and the small Ringelbach research catchment in Vosges, France as in Pappenberger et al. (2004); Freer et al. (1996); and Khu and Madsen (2005). Another popular method is the shuffled-complex evolution method discussed in Guan et al. (1992) and developed at the University of Arizona and applied to numerous watershed problems. Several recent developments in optimization are built on the basis of the Shuffled Complex Evolution (SCE) method. For instance, Gupta et al. (1998) used three objectives to calibrate 13 parameters using multiobjective complex evolution (MOCOM-UA). The procedure required 25,702 function calls and resulted in 500 Pareto solutions. Application of SCE-UA to the same problem with one objective required 5,000 to 10,000 function calls. Since computational time for problems solved with ADH can become quite large, it is not feasible to require thousands of function calls. For this study, focus was put on local methods to reduce the number of function calls, i.e., model runs.

UCODE: UCODE is a local method which performs inverse modeling using nonlinear regression developed by Poeter and Hill (1998). The inverse model is posed as a parameter-estimation problem. The nonlinear regression problem is solved by minimizing a weighted least-squares objective function with respect to the model parameters using a modified Gauss-Newton method. Forward and central differences are used to calculate sensitivities. UCODE, a universal inverse modeling program, consists of algorithms programmed in Perl and Fortran90. An advantage of UCODE is the speed of the Gauss-Newton method for well-posed problems. In instances where the observations provide adequate information about the parameters, this method is typically 10 to 100 times faster than global search methods. Other advantages include the propagation of uncertainty of all defined parameters, proven ability to perform well in many problems with substantial nonlinearity, and availability as freeware and open source. Disadvantages of this package are its inability to find a global minimum when multiple minima occur and poor performance in extremely nonlinear parts of a solution. Additionally, the code is rather cumbersome to implement requiring Fortran90 as well as Perl to run. UCODE has been applied to various groundwater applications and surface-water applications.

PEST: PEST is a local method for model-independent parameter estimation. The method is based on the Gauss-Marquardt-Levenberg method which minimizes discrepancies between model-generated numbers and corresponding field or laboratory data in a weighted least-squares sense. Advantages of PEST are that it offers many techniques from pragmatic to fully Bayesian, some techniques are numerically fast and stable, interaction through model input files makes for usability with arbitrary models, and the code is readily available. PEST is also available for parallel computing which reduces the numerical intensity of some techniques by making model runs simultaneously on a cluster of computers rather than sequentially on one computer. A disadvantage is that some techniques are inefficient when the parameters are highly correlated.

PEST offers model-independent parameter estimation with advanced predictive analysis and regularization features. More specifically, PEST is a stand-alone code that reads and manipulates model input files and reads model output files via user instructions. Though combined with ADH in this study, PEST can easily be coupled with any model by generating three input files. The Gauss-Marquardt-Levenberg method requires that a continuous relationship exist between model parameters and model output but generally finds the minimum in fewer function calls than other optimization packages. This is important when dealing with large-scale problems which are computationally intensive. Though the method is said to have a tendency to get stuck in local minima, this can be avoided by formulating an objective function which includes processed flow data as well as flows. The adjustable parameters can be bounded, which enhances the numerical stability of the parameter estimation process as the bounds are imposed. The objective function can contain numerous aspects of the system using weighting.

Once the parameter estimation package has determined a solution, the solution can then be supplied to PEST in regularization mode to obtain other sets of parameters which could also be considered to calibrate the model. The user supplies a default system condition expressed in terms of preferred values for parameters. PEST then calibrates within a preferred model-to-measurement fit tolerance defined by limiting measurement function below which the model is deemed to be calibrated. Simultaneously a regularization objective function calculated on the basis of the misfit between optimized parameter values and their user-supplied default values or relationship values is minimized.

PEST is widely used as a parameter estimation tool coupled with various simulation tools. Baginska et al. (2003) coupled PEST with Annualized Agricultural NonPoint Source (AnnAGNPS) to determine the export of nitrogen and phosphorous through nonpoint sources. Application with Surface-Water Analytical Tools (SWAT) to model snowmelt hydrology was made by Wang and Melesse (2005). Urban runoff models and watershed models such as Hydrologic Simulation Program Fortran (HSPF) have also been coupled with PEST (Doherty and Johnston 2003; Cocca et al. 2003; Ovbiebo and Kuch 1998). PEST has also been successfully applied with temperature and salinity models as in Gao and Meerick (1996). The field of groundwater has found PEST to be a useful tool with application for flow, heat transfer, and mass transfer (Keating et al. 2003; Shook and Renner 2002; Vesselinov et al. 2001 Vesselinov et al. 2002; Doherty 2003; Zyvoloski et al. 2003).

SIMULATION MODELS: The representation of physical problems with numerical models offers a number of advantages over physical model representations. Major advantages of numerical models are their relative speed and low cost. As computers have improved, so has the ability to model larger and larger physical problems using numerical models. Numerical models are used to model surface-water problems, groundwater problems, overland flow problems, transport problems, and numerous other fields. Alone, these models are useful tools for engineers. However, coupled with optimization techniques these models have the ability to test multiple scenarios in an intelligent manner to provide a better solution. One numerical model that provides a mechanism for modeling groundwater, surface water, overland flow, and transport is ADH (Howington et al. 2003; Berger and Schmidt 2004).

ADH: ADH is a finite element model that uses linear simplex elements to represent the physical problem and offers mesh refinement and coarsening in both serial and parallel computing platforms. ADH was developed to address the environmental concerns of the Department of Defense in estuaries, coastal regions, river basins, reservoirs, and groundwater. The basic philosophy behind ADH is to offer a centralized computational engine which includes such things as solvers, preconditioners, finite element utilities, and input/output that can be reused to solve multiple hydrologic components separately or in a coupled fashion. The parallel computing capabilities are the Message-Passing model (MPI) and are necessary to run the large-scale problems that are of interest to the DOD. The goal of this research is to determine efficient techniques for parameter estimation for large-scale, multiphysics problems like the ones solved by ADH.

APPLICATIONS: In this study, PEST was used with ADH and applied to several physical problems. PEST modifies ADH input files by adjusting the model parameters and reads ADH output files to determine the resulting objective function. Figure 1 gives a simple flow chart of how the two packages interact. Initially, the coupling was tested on the Theis problem, a well-known problem in groundwater flow with an analytical solution. This problem consists of a cube with one pumping well located in the center (Figure 2). The parameters being varied are hydraulic conductivities in the x, y, and z directions for the subsurface material present. The mesh consists of 9,483 nodes and 37,440 elements. For the PEST runs, nine observation points are used with head values. Each model run requires 4.6 min of Central Processing Unit (CPU) time. PEST determined the conductivity values with six optimization cycles requiring 20 ADH runs. Figure 3 shows the models calls related to the optimization cycle. A convergence history is given in Figure 4 where the conductivity values are shown related to the optimization cycle. Finally, Figure 5 shows the history of the objective function as it relates to the optimization cycles. Notice how the objective function was reduced from an initial value of over 650 to a value near 0. Since the objective function is a least squares fit of the computed heads to analytical values, a small objective function indicates a good fit between the computed and known values. Using an initial conductivity value of 0.4 for all directions, the model determined conductivity in the x-direction to be 0.500462 with a 95 percent confidence interval of 0.499112 to 0.501812, the conductivity in the y-direction to be 0.499490 with a 95 percent confidence interval of 0.498167 to 0.500814, and a conductivity value in the z-direction of 0.499969 with a 95 percent confidence interval of 0.459354 to 0.540585. These conductivities compared well with the conductivities prescribed for the analytical solution which were 0.5 ft/day in each principle direction. In terms of conductivity values, the 95 percent confidence intervals are small.

PEST was also used with ADH to solve a surface-water problem. This problem is a section of the Mississippi River called Pool 8 (Figure 6). The parameter being varied is the Manning's roughness coefficient for the single surface material present. The mesh consists of 10,352 nodes and 17,103 elements. For the PEST runs, depths are specified at 98 observation points. Each model run requires 28 min of CPU time. PEST determined the conductivity values with six optimization cycles requiring nine model runs. Figure 7 shows the model calls related to the optimization cycle. A convergence history is given in Figure 8 where the conductivity values are shown related to the optimization cycle. Finally, Figure 9 shows the history of the objective function as it relates to the optimization cycles. Notice how the objective function was reduced

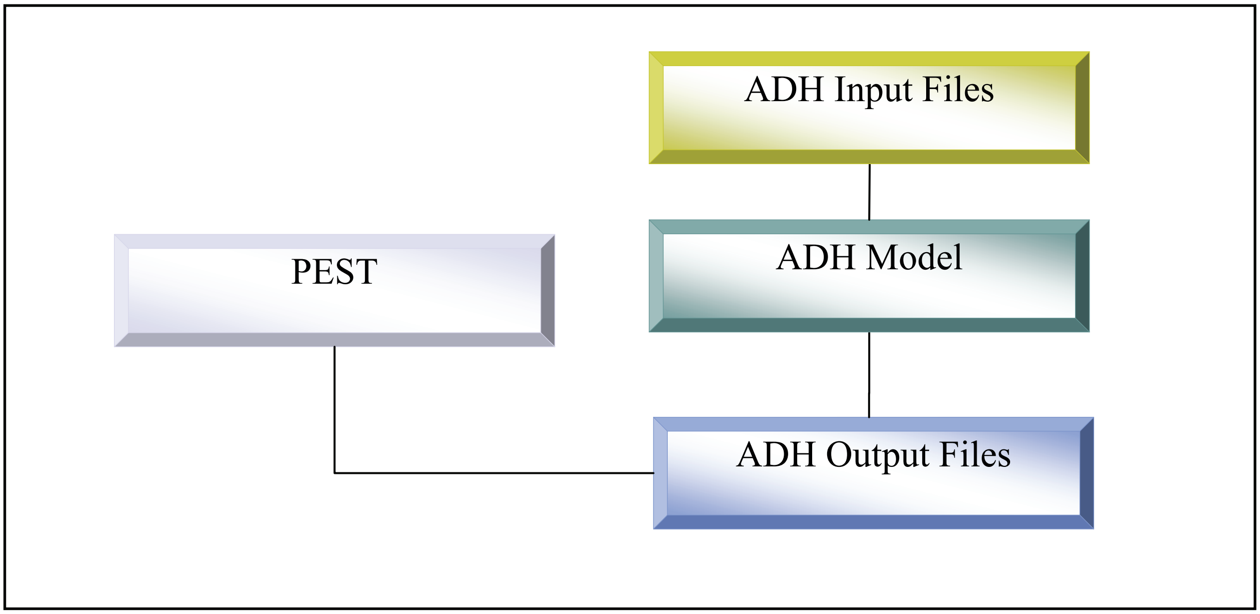

Figure 1. Flow chart of PEST and ADH interaction.

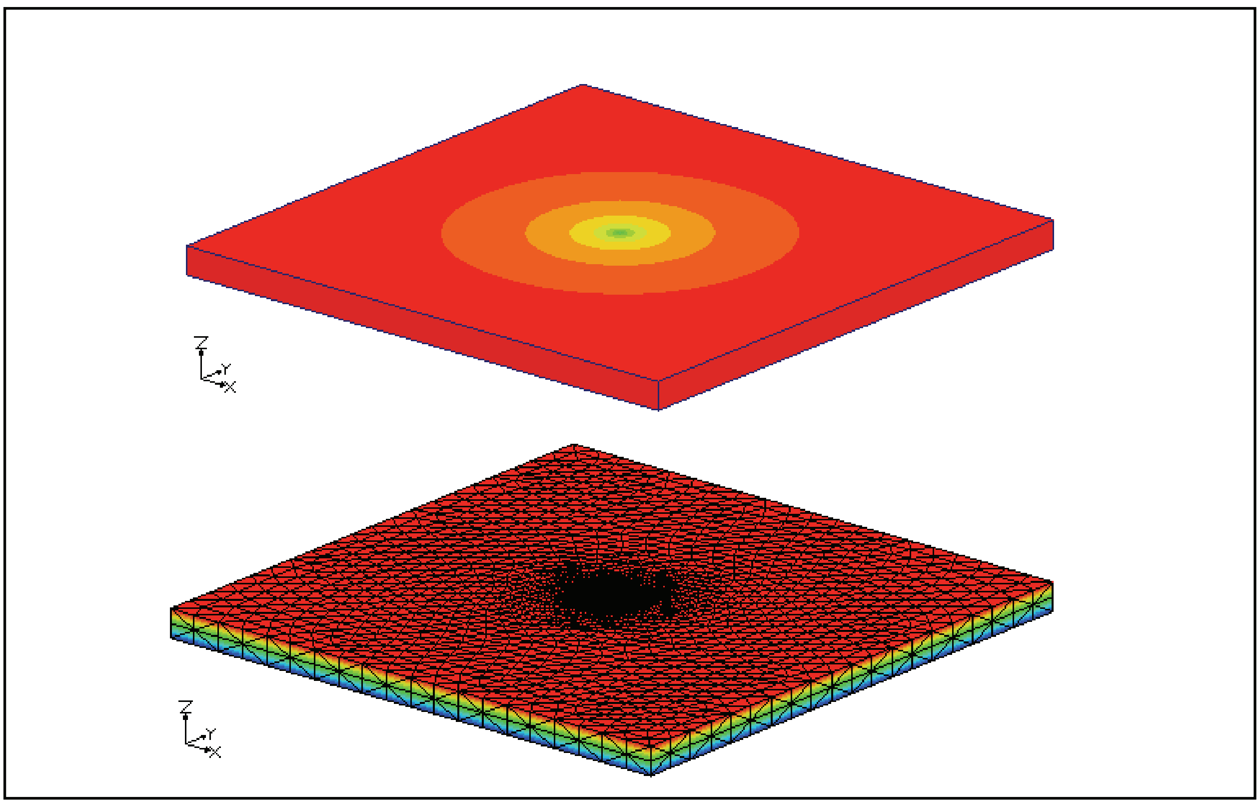

Figure 2. The Theis problem mesh and head values.

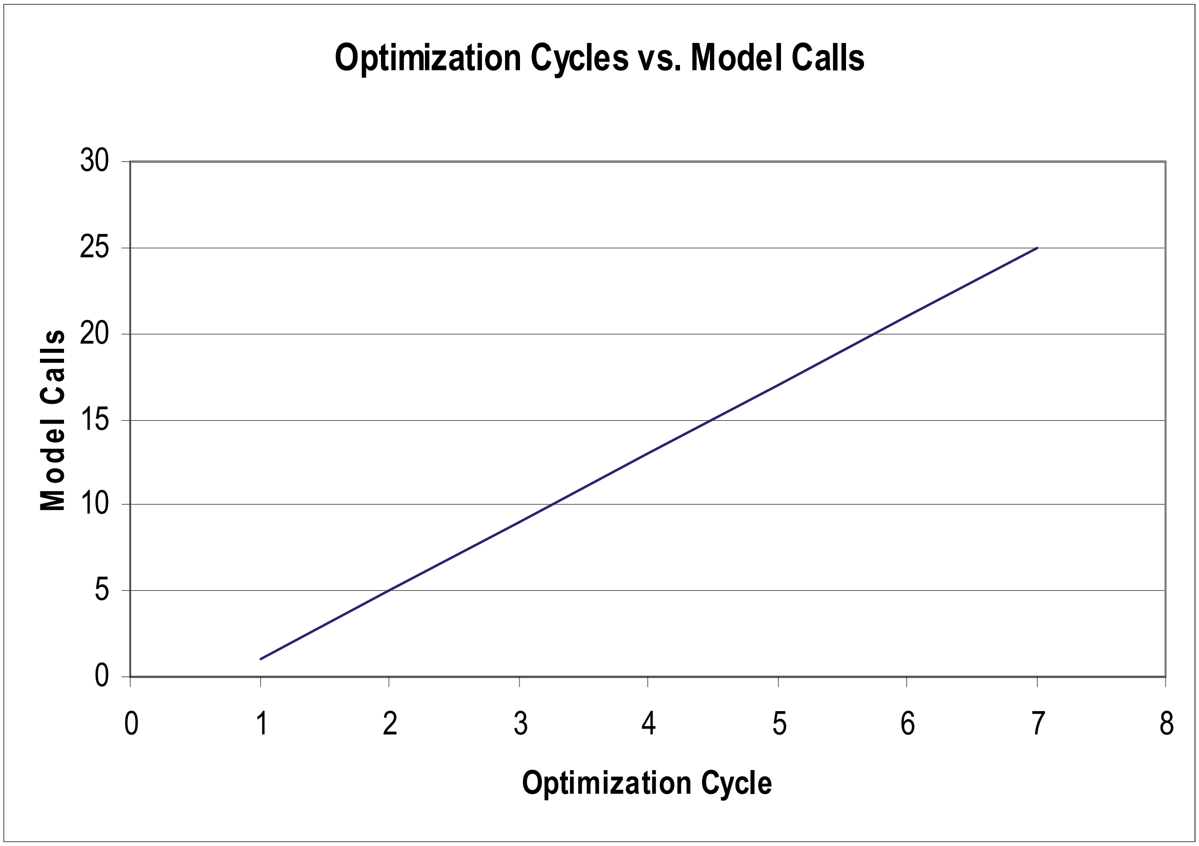

Figure 3. Optimization cycle versus model calls for Theis problem.

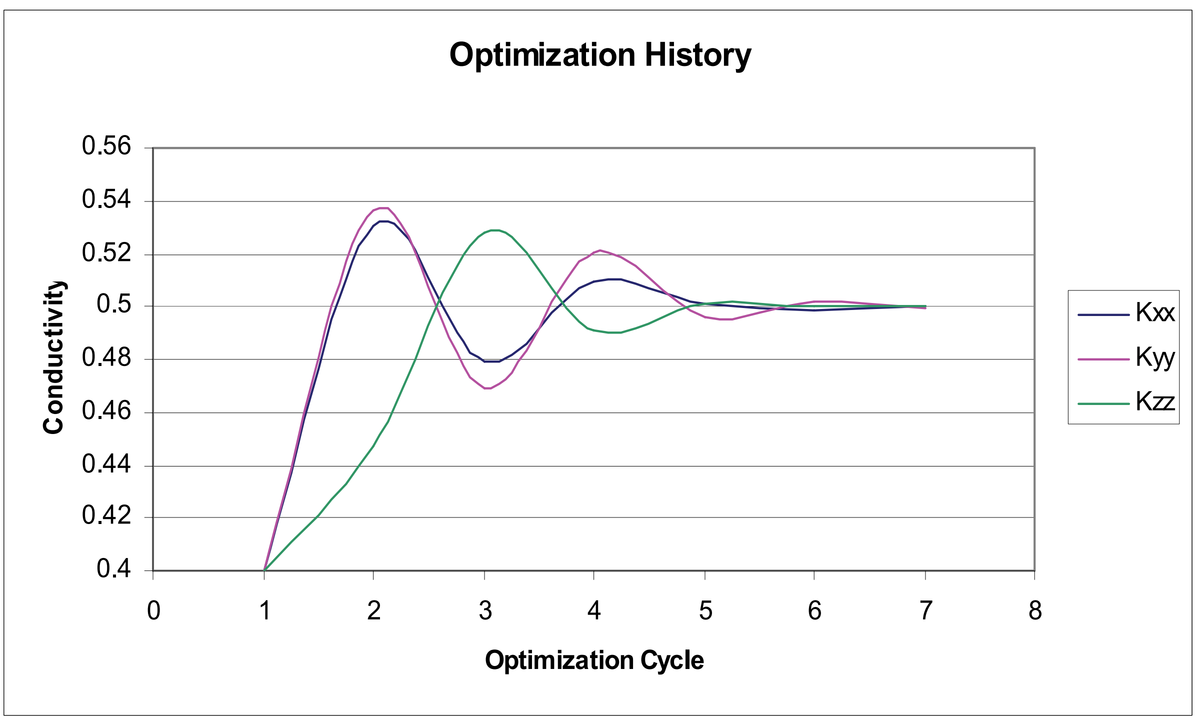

Figure 4. Optimization cycles versus conductivities for Theis problem.

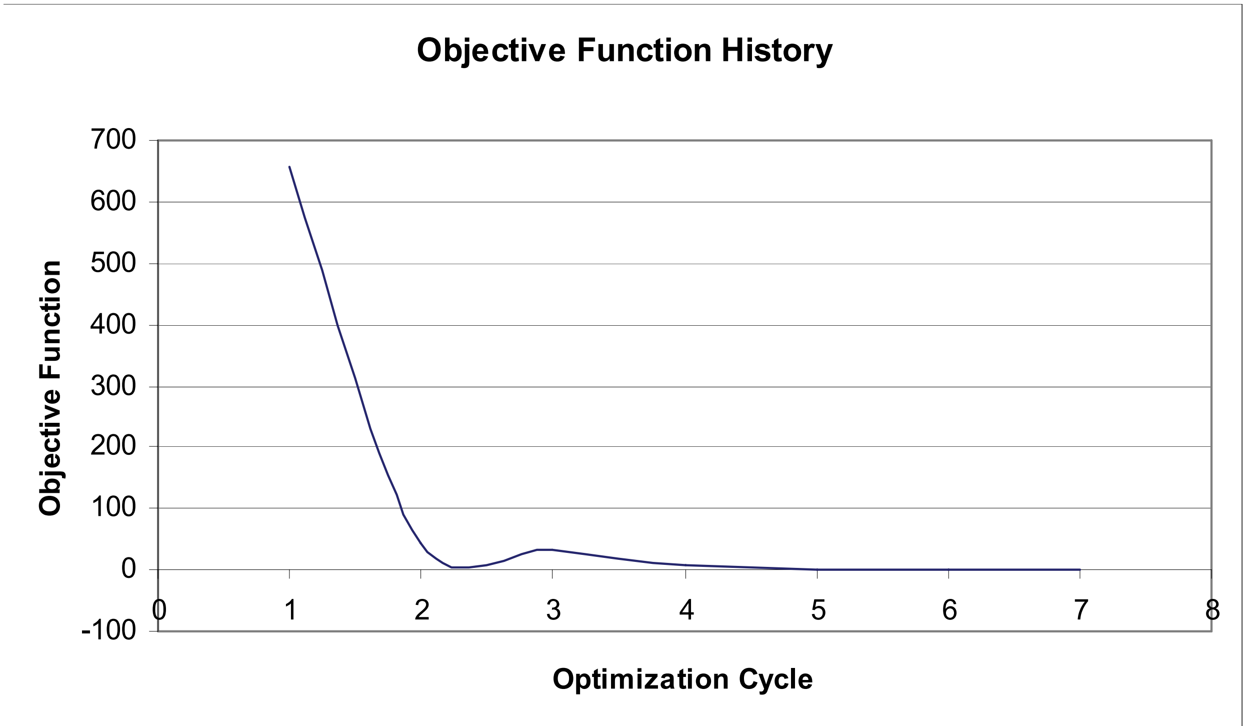

Figure 5. Objective function history for Theis problem.

from an initial value of over 12,000 to a value near 0, again indicating a good fit between model and data. Since the objective function is a least squares fit of the computed depths to the depths given from field data, the objective function obtained shows that the computed head values are near if not identical to the field data. The model determined the roughness value to be 0.02502 with a 95 percent confidence interval of 0.02500351 to 0.02503649. The confidence interval indicates that the value for roughness is within a small range.

SUMMARY: Optimization techniques coupled with numerical models provides an effective mechanism for estimating input parameters to obtain a better fit to field data. Though a number of techniques exist, few have the ability to handle large-scale, multiphysics problems like the ones solved with ADH in a computationally-efficient manner. PEST is a gradient-based estimation package that has been extended to deal with many of the difficulties encountered with gradient-based methods. The test cases discussed in this work gave good preliminary results that indicated that PEST may be used for parameter estimation with ADH for small-scale, multiphysics problems. However, as the number of input parameters and model size becomes larger and larger, adjustments will be required to the optimization package to reduce the number of model calls or reduce the computational time required for each model call. For this reason, the parallel version of PEST will be tested to allow for simultaneous parallel model calls to reduce the overall computational time. Also, techniques that include stochastic approximation and simultaneous perturbations to determine the derivatives will be applied.

Figure 6. Pool 8 mesh and bed elevations.

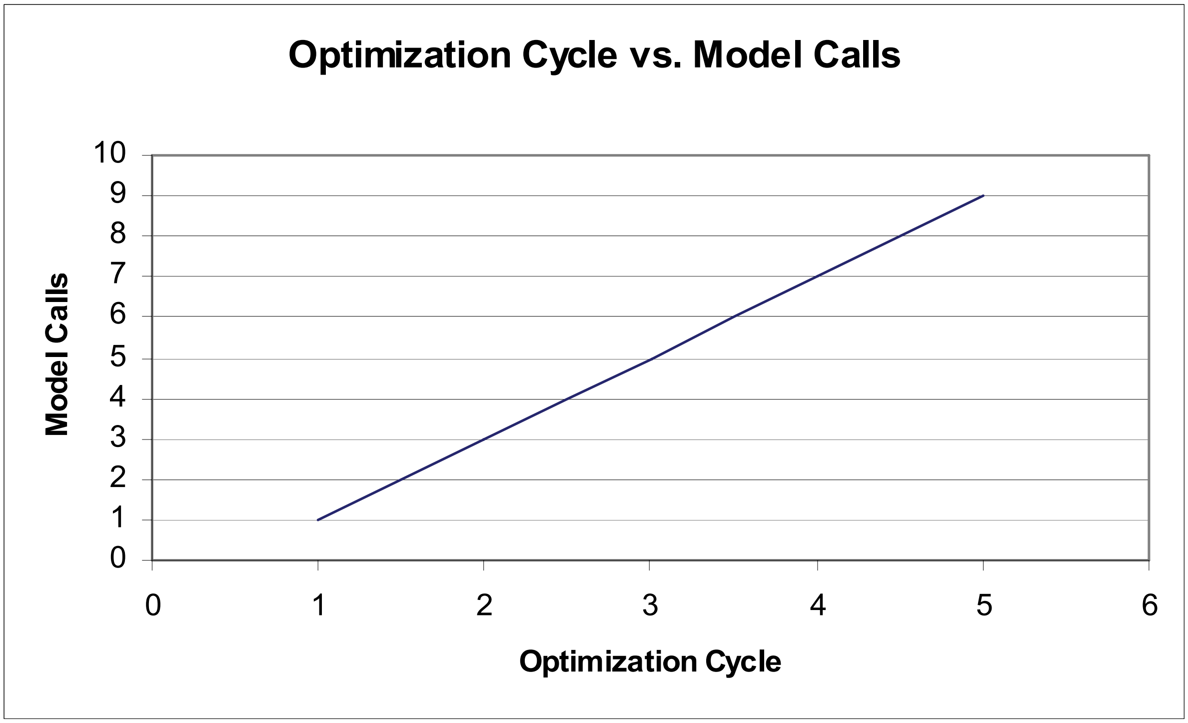

Figure 7. Optimization cycles versus model calls for Pool 8.

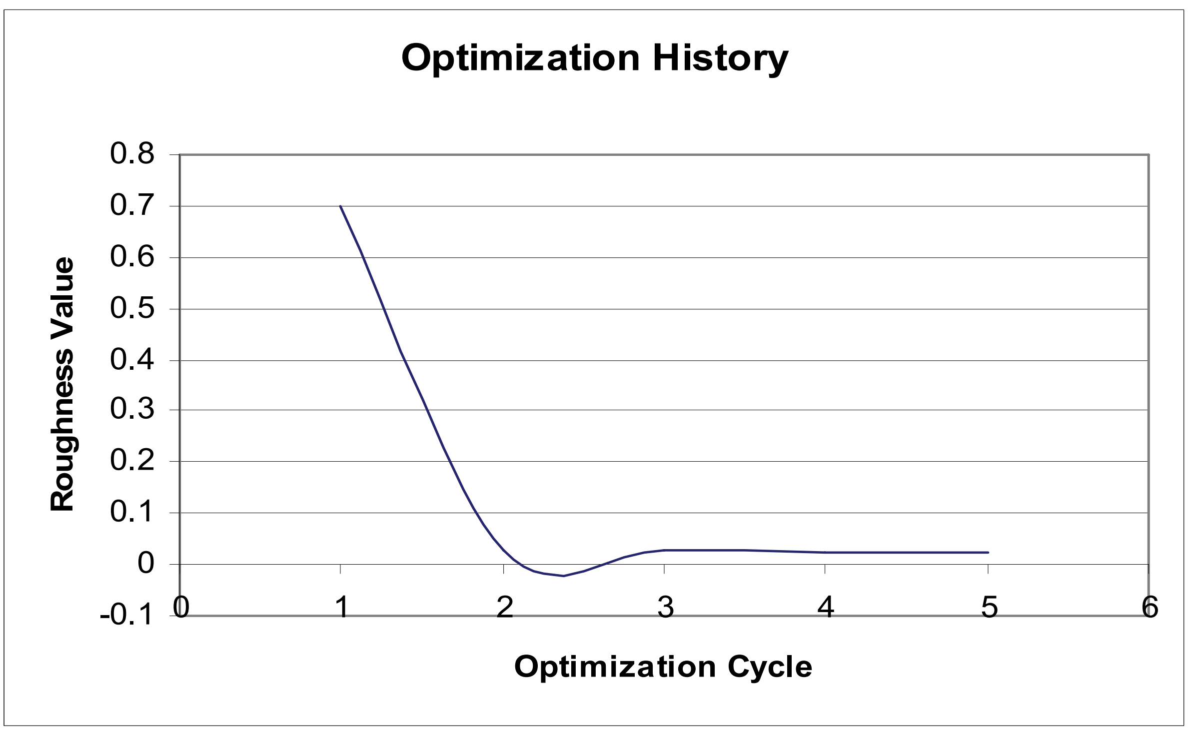

Figure 8. Roughness value history for Pool 8.

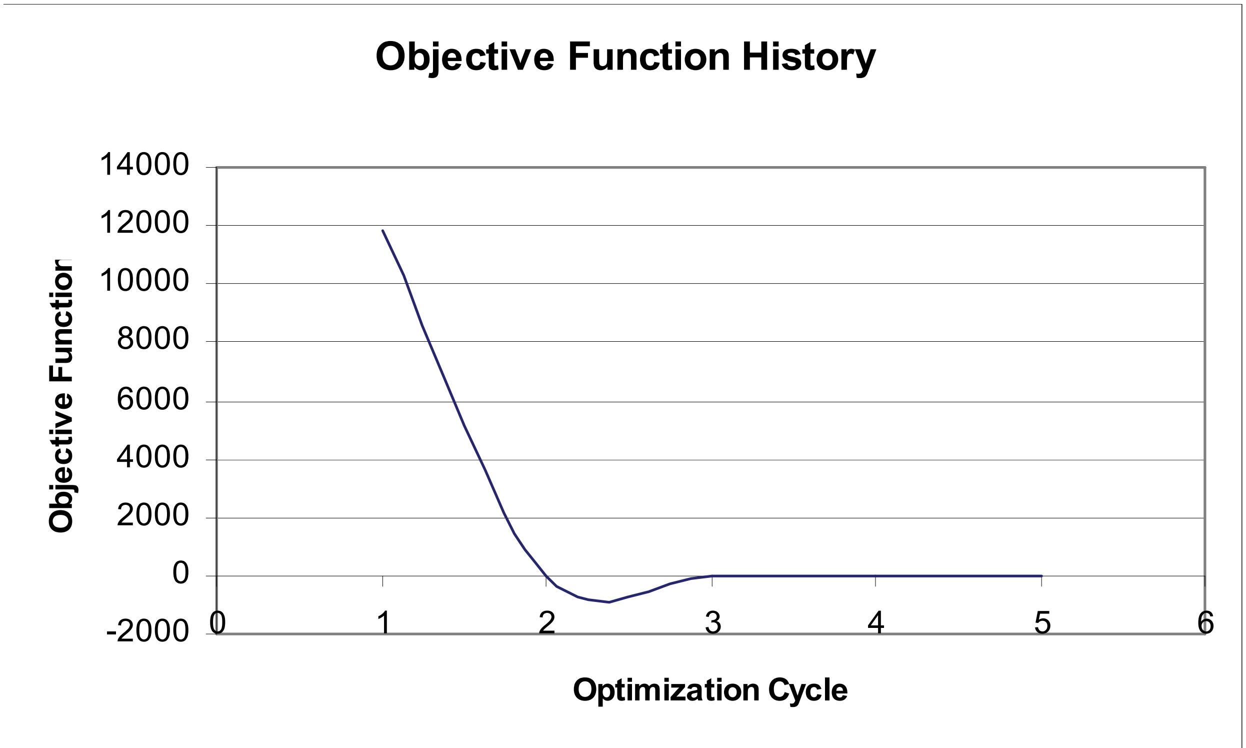

Figure 9. Objective function history for Pool 8.

ADDITIONAL INFORMATION: This technical note was prepared by Jackie P. Hallberg, Coastal and Hydraulics Laboratory, U.S. Army Engineer Research and Development Center, 3909 Halls Ferry Road, Vicksburg, MS 39180. The study was conducted as an activity of the Watershed Hydrology Simulation work unit of the System-Wide Water Resources Program. For information on SWWRP, please consult https://swwrp.usace.army.mil/ or contact the Program Manager, Dr. Steven L. Ashby at Steven.L.Ashby@erdc.usace.army.mil. This technical note should be cited as follows:

> Hallberg, J. P. 2006. *Parameter estimation tools for hydrologic and hydraulic simulations.* ERDC TN-SWWRP-06-11. Vicksburg, MS: U.S. Army Engineer Research and Development Center. https://swwrp.usace.army.mil/

REFERENCES

Baginska, B., W. Milne-Home, and P. S. Cornish. 2003. Modelling nutrient transport in Currency Creek, NSW with AnnAGNPS and PEST, *Environmental Modelling and Software* 18(8-9): 801-808.

Berger, R. C., and J. H. Schmidt. 2004. *Guidelines for Solving Three Dimensional Navier-Stokes Problems with the Adaptive Hydraulics (ADH) Modeling System*. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Beven, K. J., and J. Freer. 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems. *J. Hydrology* 249, 11-29.

Cocca, P. A., J. Doherty, and J. L. Kittle Jr. 2003. Hydrologic Calibration strategies for the HSPF Watershed Model: Identifying Effective Objective Functions for Use with the Parameter Estimation (PEST) Program. *World Water and Environmental Resources Congress* 2429-2438.

Doherty, J. 2003. Ground water model calibration using pilot points and regularization, *Ground Water* 41(2): 170-177.

Doherty, J., and J. M. Johnston. 2003. Methodologies for calibration and predictive analysis of a watershed model, *J. American Water Resources Association* 39(2): 251-265.

Fletcher, R., *Practical Methods of Optimization*, 2nd Edition, John Wiley and Sons, 1987.

Freer, J., K. Beven, and B. Ambroise. 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resour. Res*. 32(7), 2161-2173.

Gao, J., and N. P. Merrick. 1996. Simulation of temperature and salinity in a fully mixed pond. *Environmental Software*, 11(1-3), *Regional Development and Environmental* Change, 173-178.

Guan, Q., S. Sorooshian, and V. Gupta. 1992 Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models., *Water Resour. Res*. 28(4), 1015-1031.

Gupta, H. V., S. Sorooshian, and P. O. Yapo., 1998. Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information. *Water Resour. Res*. 34(4), 751-763.

Gutierrez-Magness, A. L., and R. H. McCuen., 2005. Effect of flow proportions on HSPF model calibration accuracy. *J. of Hydrologic Engineering* 10(5): 343-352.

Howington, S. E., R. C. Berger, E. W. Jenkins, J. P. Hallberg, K. R. Kavanagh, and J. H. Schmidt. 2003. *User's Manual for Groundwater/Surface Water Interaction Using ADH*. Vicksburg, MS: U.S. Army Engineer Research and Development Center.

Keating, E. H., V. V. Vesselinow, E. Kwicklis, and Z. Lu. 2003. Coupling basin-and site-scale inverse models of the Espanola aquifer., *Ground Water* 41(2): 200-211.

Kelley, C. T. 1999. *Iterative methods in optimization*. Society for Industrial and Applied Mathematics.

Khu, S. T., and H. Madsen. 2005. Mutiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration, *Water Resour. Res.* 41, W03004, doi: 10.1029/2004WR003041.

Nocedal, J., and S. J. Wright. 1999. *Numerical optimization*. Springer Series in Operations Research, Springer Science.

Onwubiko, C. 2000. *Introduction to engineering design optimization*. Prentice-Hall, Inc.

Ovbiebo, T., and A. W. Kuch. 1998. Non-linear parameter estimation of an urban runoff model using XP-SWMM32 and pest. *Proceedings of the Annual Water Resources Planning and Management Conference, Water Resources and the Urban Environment*, 247-252.

Pappenberger, F., K. Beven, M. Horritt, and S. Blazkova. 2004. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *J. of Hydrology* 302, 46-69.

Poeter, E. P., and M. C. Hill. 1998. Documentation of UCODE, a computer code for universal inverse modeling: U.S. Geological Survey Water-Resources Investigations Report 98-4080.

Shook, G. M., and J. L. Renner. 2002. An Inverse Model for TETRAD: Preliminary Results, Geothermal Resources Council: 2002 Annual Meeting, Sep 22-25 2002, Reno, NV, United States.

Vanderplaats, G. N. 2001. *Numerical optimization techniques for engineering design*, 3rd ed. Colorado Springs, CO: Vanderplaats Research and Development, Inc.

Vesselinov, V. V., S. P. Neuman, and W. A. Illman. 2001. Three-dimensional numerical inversion of pneumatic cross-hole tests in unsaturated fractured tuff 1. Methodology and borehole effects. *Water Resour. Res*. 37(12): 3001-3017.

Vesselinov, V. V., E. H. Keating, and G. Zyvoloski. 2002. Analysis of model sensitivity and predictive uncertainty of capture zones in the Espanola basin regional aquifer, northern New Mexico. *Calibration and reliability in groundwater modeling*, 508-514, *IAHS-AISH Publication*, n 277.

Vrugt, J. A., C. G. H. Diks, H. V. Gupta, W. Bouten, and J. M. Verstraten. 2005. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res*. 41, W01017, doi: 10.1029/2004WR003059.

Wang, X., and A. M. Melesse. 2005. Evaluation of the SWAT model's snowmelt hydrology in a northwestern Minnesota watershed, *Transactions of the American Society of Agricultural Engineers* 48(4): 1359-1376.

Zyvoloski, G., E. Kwicklis, A. A. Eddebbarh, B. Arnold, C. Faunt, and B. A. Robinson. 2003. *Journal of Contaminant Hydrology* 62-63: 731-750.